



Posudek oponenta závěrečné práce

Oponent práce:	Ing. Magda Friedjungová, Ph.D.
Student:	Vít Kalianko
Název práce:	Doporučovací systém pro volbu předmětu na základě podobnosti studentů
Obor / specializace:	Znalostní inženýrství
Vytvořeno dne:	6. června 2023

Hodnotící kritéria

1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Zadání je dle mého splněno, i když by v něm bylo vhodné definovat, co je myšleno "doporučovacím systémem".

2. Písemná část práce

75 / 100 (C)

Práce je psána v češtině a je logicky uspořádána. Teoretickou část bych si dovedla představit stručnější s apelem na pokročilejší metody nebo zajímavosti o těch základních (např. rozsáhlejší rešerše, kde se kNN používá). Stejně tak by jednotlivé kapitoly nemusely být tolik členité. Popis praktické části je o poznání lepší, nicméně celkově je práce pojata vypravěčským způsobem, kdy místy postrádám fakta a je pro mě čtení o to náročnější. Experimenty jsou víceméně dobře popsány, i když bychom mohli vést diskuzi o vhodnosti nastavení hyperparametrů a možnosti srovnání s jinými pracemi (které je v práci opravdu pouze orientační). Níže uvádím pár výtek.

- Práce obsahuje několik překlepů. V textu se občas objeví kombinace češtiny a anglického pojmu (např. "v rámci information retrieval", "accuracy" apod.)
- Seznam zkratk není kompletní, chybí např. KOS. Resp. některé zkratky jsou zavedeny pouze v textu (např. PP, PV apod.), jiné jsou pouze v seznamu zkratk.
- Úvodní text si dovedu představit méně zabarvený a podložený statistikami (např. "obrovská nabídka předmětů" si říká o číslo).
- Na můj vkus je text poněkud lidový/vypravěčský a bohatý na různá hodnocení ze stran studenta bez uvedení referencí.
- Část 2.1 Vymezení pojmů bych vynechala - jedná se pouze o opis pro náš obor zřejmého.

- Chápu, že je práce postavena především na kNN, a proto jsou mu v teoretické části věnovány 4 stránky. Nicméně je to trochu přepis přednášky z BI-VZD/BI-ML1.
- V části 6.4 mi chybí konkrétní čísla. Nerozumím, proč je nejdříve uvažována testovací množina a z ní je vytvořena trénovací, která by měla být typicky větší, tedy opačný přístup by zabral méně času. V sekci 8.1 je to popsáno trochu lépe.
- Str 49., poznámka pod čarou by měla být až pod tabulkou.

3. Nepísemná část, přílohy

92 /100 (A)

Student pro svoji práci zvolil přiměřené technologie a odevzdal několik okomentovaných a spustitelných jupyter notebooků včetně potřebných dat. V zadání práce by bylo vhodné definovat, co je myšleno "doporučovacím systémem", lehce toto spojení zavání aplikací, kterou student nerealizoval (po doptání ani neměl). Velkou část práce a času zabírá příprava dat, se kterou se student vypořádal statečně. Navržené řešení je funkční a je za ním vidět studentova píle a snaha.

4. Hodnocení výsledků, jejich využitelnost

78 /100 (C)

Student odvedl dobrou práci, kdy dosáhl uspokojivých výsledků, resp. i lepších, než v pracech, na které navazoval. Výsledky by bylo vhodné ještě nasadit do praxe nebo je alespoň na jednom novém semestru otestovat. Na práci lze dále navázat, student zmiňuje několik možností. Za nejpřínivější považuji doporučení pro magisterské studenty, které mohlo být také realizováno. Po mentální stránce se jedná o nelehkou úlohu, kdy existuje několik úhlů, jak s daty pracovat, jaký dataset pro doporučení sestavit (student musel učinit několik rozhodnutí, např. vůči různým akreditacím), stejně tak je několik možností, jaké doporučovací metody experimentálně ověřit. V práci nejsou pokryty všechny možnosti (např. nejsou nakonec zohledněna data z Ankety), což nebylo ani cílem, zároveň je ale dle mého zvolena nepřiliš složitá cesta.

Celkové hodnocení

79 /100 (C)

Práci navrhuji hodnotit klasifikačním stupněm C, nicméně balancuji na hranici, kdy je mé hodnocení možná přísnější. Student splnil zadání, a přestože danou úlohu řeší jednodušším algoritmem, jedná se o komplexní řešení, které může být dále využito či rozšířeno. Bylo by pěkné, pokud by student chtěl práci nasadit do praxe, např. v rámci dalšího studia.

Otázky k obhajobě

1. Mohl byste prosím popsat, jak byla vytvořena trénovací a testovací data a jejich přesný počet?
2. Co si myslíte o robustnosti vašeho řešení? Neuvažoval jste algoritmus náhodného lesa, případně jiné varianty rozhodovacích stromů, které by mohly lépe reagovat na odlehle hodnoty v datech? Např. jiné chování při covidu, různé strategie zápisu kreditů napříč semestry apod.

Instrukce

Splnění zadání

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Nepísemná část, přílohy

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Celkové hodnocení

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.